Title: Identification of amino acids within norovirus polymerase involved in RNA binding and viral replication.

Kang Rok Han[1, 2], Bader Y Alhatlani[3], Sunyoung Cho[2], Ji-Hye Lee[1], Myra Hosmillo[3], Ian G Goodfellow[3], Kyung Hyun Kim[1]*, and Jai Myung Yang[2]*

[1]Department of Biotechnology & Bioinformatics, Korea University, Sejong 339-700, South Korea

[2]Department of Life Science, Sogang University, Seoul 121-742, South Korea

[3]Division of Virology, Department of Pathology, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2-2QQ, United Kingdom

*Correspondence to Kyung Hyun Kim, khkim@korea.ac.kr and Jai Myung Yang, jaimyang@sogang.ac.kr

Running title: Role of basic amino acids at the catalytic site of MNV-1 NS7

Keywords: norovirus, RNA binding, RNA replication

Subject category: Animal – Positive strand RNA viruses

Words in summary: 145

Words in text: 2172

Summary

Until recently, molecular studies on human norovirus (HuNoV), a major causative agent of gastroenteritis, have been hampered by the lack of an efficient cell culture system. Murine norovirus-1 (MNV-1) has served as a surrogate model system for norovirus (NoV) research, due to the availability of robust cell culture systems and reverse genetics. To identify amino acids involved in RNA synthesis by the viral RNA-dependent RNA polymerase (NS7), we constructed NS7 mutants in which basic amino acids surrounding the catalytic site were substituted with alanine. Electrophoretic mobility shift assay (EMSA) revealed that these residues are important for RNA binding, particularly R^{396}. Furthermore, in vitro RNA synthesis and reverse genetics was used to identify conserved amino acids essential for RNA synthesis and viral replication. These results provide additional functional insights into highly conserved amino acids in NS& and provide potential methods of rational attenuation of norovirus replication.

Human noroviruses (HuNoV) are considered a major cause of nonbacterial acute gastroenteritis worldwide (Green, 2007). Studies on the molecular mechanisms of HuNoV replication and its pathogenic properties have been hampered by the lack of an efficient virus propagation system in cell culture and an animal model. The recent development of two culture systems for human norovirus provides additional experimental system with which to dissect the molecular mechanisms involved in norovirus genome replication (Jones et al., 2014; Ettayebi et al., 2016). Since 2003, however, murine norovirus-1 (MNV-1) has been used as a the most robust experimental model due to the availability of a robust cell culture system, reverse genetics and small animal model (Karst et al., 2003; Wobus et al., 2004), and has served as a surrogate model system to study norovirus (NoV) replication and some aspects of viral pathogenicity (Karst et al., 2014).

The MNV-1 genome is a positive-sense single-stranded RNA of approximately 7.4 kb. A virus encoded protein VPg is covalently attached to the 5' end of the RNA (Olspert et al., 2016), with a poly(A) tail at the 3' end (Karst et al., 2003; Daughenbaugh et al., 2003). Of the four open reading frames (ORFs) within the MNV genome, ORF1 encodes a polyprotein that is cotranslationally processed into six functional proteins by a virally encoded protease (NS6): NS1-2 (N-term), NS3 (NTPase), NS4, NS5 (VPg), NS6 (Pro), and NS7 (RdRp, Pol) (Sosnovtsev et al., 2006). In addition to full-length genomic RNA (gRNA), a 2.5-kb-long subgenomic RNA (sgRNA) carrying ORF2 and ORF3 is also produced in MNV-1–infected cells (Wobus et al., 2004). ORF2 and ORF3 encode the major capsid (VP1) and minor capsid (VP2) proteins, respectively. ORF4, encoding the VF1 protein, functions to antagonize the innate immune response to infection (McFadden et al., 2011).

Upon infection with members of the *Caliciviridae* family, the positive-sense full-length genomic RNA is expected to play a dual role. The RNA is used as a template for translation to produce nonstructural proteins and used as a template to generate negative-sense RNA, which will serve as a template to generate gRNA and sgRNA (Thorne and Goodfellow, 2014). A virally encoded RNA-dependent RNA polymerase (RdRp) and its uncleaved proteinase-polymerase precursor are involved in the synthesis of viral RNAs (Fukushi et al., 2004; Rohayem et al., 2006; Medvedev et al., 2017). Sequence comparison and structural analysis of the rabbit hemorrhagic disease virus RHDV RdRp identified six motifs within the palm domain, including the active site YGDD-containing motif C and motif A (Fig 1a) (Vázquez et al., 2000; Ng et al., 2002). In addition, crystallographic studies

showed that the overall structure of the RHDV (Ng et al., 2002), NV (Ng et al., 2004), sapovirus (SV) (Fullerton et al., 2007), and MNV (Lee et al., 2011) RdRps are similar to those of poliovirus (Hansen et al., 1997), hepatitis C virus (HCV) (Bressanelli et al., 1999), and consist of a cupped right hand with palm, fingers, and thumb domains (Fig. 2a). The N-terminal domain of NV, RHDV, MNV, and SV RdRp connect the fingers and thumb domains, whereas the C-terminal domain is located within the active site cleft and potentially interferes with primer and/or template RNA during viral RNA synthesis.

We have previously demonstrated that recombinant MNV-1 NS7 expressed and purified from *Escherichia coli* is functionally active, possesses RNA synthesis activity and retains a closed right hand structure, with palm, fingers, and thumb domains (Lee et al., 2011; Han et al., 2010). To identify specific amino acid(s) involved in RNA binding and RNA synthesis activities conserved across the *Caliciviridae* family, the sequences of RdRps from NV, RHDV and MNV-1 were aligned (Fig. 1a). We assumed that the basic amino acids residing within or adjacent to the active site of NS7 are likely involved in interacting with template RNA. Amino acid sequence alignment revealed that MNV-1 residues K^{169}, K^{210}, K^{183}, K^{184}, R^{185}, R^{395}, R^{396}, and K^{422} are well conserved between these viruses. Studies of the crystal structure of RHDV RdRp-primer-template triplex indicated that residues K^{173}, K^{213}, K^{403}, and R^{404} are involved in the interaction with the RNA template, and that amino acid R^{188} interacts with the nucleotide (Ng et al., 2002).

Residues K^{169}, R^{185}, K^{210}, R^{395}, and R^{396} of the MNV-1 NS7 correspond to K^{173}, R^{188}, K^{213}, K^{403}, and R^{404} of RHDV RdRp. In addition to these amino acids, a number of basic amino acids (K^{183}, K^{184}, R^{185}, and K^{422}) at the active site were found to be located near the RNA template when we superimposed our crystal structure of the MNV-1 NS7 on structure polymerase-RNA complex of NV (PDB ID 3BSO) (Fig. 2). These results suggested that these amino acid residues could be involved in the MNV-1 NS7-RNA interface.

To examine the functional roles of the residues K^{169}, K^{183}, K^{184}, R^{185}, K^{210}, R^{395}, R^{396}, and K^{422}, seven alanine substituted mutants, K169A, K210A, K422A, KKRAAA (K183A/K184A/R185A), RRAA (R395A/R396A), R395A, and R396A NS7s were constructed by site directed mutagenesis. Wild type and alanine-substituted mutants of NS7 were expressed in *E. coli* as N-terminal His-tagged proteins, which were purified by affinity chromatography as described previously (Han et al., 2010). Analysis of

92 the recombinant proteins by SDS-PAGE indicated that NS7s were purified to greater than 90%
93 homogeneity (Fig. 1b).

94 To examine the impact of the mutation on the ability of NS7 to interact with RNA, EMSA assays
95 were conducted with the Ala substitution mutant NS7s. A ^{32}P-labeled viral RNA probe was synthesized
96 by in vitro transcription using the RiboMAX RNA production system (Promega, Wisconsin, USA)
97 followed by purification with the RNeasy Mini Kit (Qiagen, California, USA) according to the
98 manufacturer's instructions. EMSA reactions contained 70 nM ^{32}P-labeled RNA probe, 1, 2, 4, and 6
99 μM of wild type or mutant NS7, 0.67 mg/ml yeast tRNA (Ambion, California, USA), 50 mM HEPES (pH
100 7.4), 25 mM KCl, 2.5 mM $MgCl_2$, 1 mM DTT, and 4% glycerol. Reactions were incubated for 10 min at
101 30°C prior to separation on 4% acrylamide gels containing 0.5X TBE and 5% glycerol. Gels were dried
102 and ^{32}P-labeled RNA was visualized using a BAS-1500 phosphorimager (Fuji Film, Tokyo, Japan). To
103 quantify complex formation, bands were analyzed using the Multi Gauge software (Fuji Film, Tokyo,
104 Japan).

105 Compared to wild type NS7, all mutants exhibited decreased RNA-binding capability to varying
106 degrees (Fig. 3). In the presence of 2 μM NS7, ~60% of the input RNA was associated with the wild
107 type NS7 in an RNA-protein complex (Fig. 3b), however in the mutants could be divided into three
108 broad groups based on the extent of RNA-protein complex formation at 2 μM NS7. The first group
109 (KKRAAA, RRAA and R396A) exhibited the greatest reduction in RNA-protein complex formation,
110 implying that these amino acid residues might play a functional role in RNA binding. The mutants
111 K210A and K422A were also significantly impaired, showing an ~60% reduction in RNA binding
112 capacity at 2 μM NS7 when compared to the wild type NS7. The third group (R395A and K169A)
113 exhibited less detrimental effects with an ~40% reduction in binding. The single amino acid alteration
114 R396A mutant exhibited a similar level of RNA-protein complex formation as with KKRAAA or RRAA,
115 whereas mutation of the adjacent residue, R^{395}, had a much-reduced impact on RNA-binding capability.
116 Further analysis of positions K^{183}, K^{184}, and R^{185} by modeling of the MNV structure on the NV RNA
117 polymerase-RNA complex, indicated that R^{185} most likely interacts with nucleotide instead of RNA (Fig.
118 2), similarly to R^{188} in the RHDV RdRp (Ng et al., 2002). Taken together, these data indicated that the
119 amino acids R^{396}, K^{183}, K^{184} and R^{185} of NS7 play a potential role in the interaction with RNA. The
120 observation that amino acid R^{404} of RHDV RdRp, equivalent to R^{396} in the MNV NS7, interacts with the

primer-template duplex also supports this hypothesis.

Since binding of template RNA to the active site of the NS7 is required for RNA synthesis, we measured the impact of the NS7 mutation on the ability of recombinant polymerase to use in vitro transcribed MNV-1 sgRNA as a template for RNA synthesis. The assay was performed as described earlier (Han et al., 2010) with minor modifications. Twenty microliters of reaction mixture containing 50 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 10 mM DTT, 1 µg sgRNA, 25 µM each NTP, 2.5 µCi [α-^{32}P]GTP (3000 Ci/mmol, 10 mCi/ml), and 0.9 µg wild type or mutant NS7s was incubated at 37°C for 30 min. The reaction was stopped by adding an equal volume of 200 mM EDTA (pH 8.0), then, 8 µl of reaction mixture was spotted onto DE81 filter paper (Fisher Scientific, Pittsburgh, USA). The filter paper was dried at room temperature for 10 min, washed three times with 2 ml 2X SSC solution for 10 min, dehydrated with 2 ml absolute ethanol, and dried at 80°C. The radioactivity of incorporated [α-^{32}P]GMP was measured with a liquid Wallac 1407 scintillation counter (Wallac, Turku, Finland). As expected, all NS7 mutants exhibited decreased RNA synthesis activities. KKRAAA, RRAA and R396A, which showed the least affinity to RNA, also showed 6%, 7%, and 12% of wild type RNA polymerase activity, and K210A, K422A, R395A, and K169A showed 42%, 25%, 35% and 20%, respectively (Fig. 3c). The extent to which RNA synthesis was impaired correlated well with the RNA binding capability determined by EMSA, with the exception of K169A.

To examine whether the observed effects on NS7 mutations on RNA binding also impact MNV replication in cell culture, we introduced mutations into the MNV-1 cDNA clone, and the effect of the mutations on virus viability examined by reverse genetics as described (Chaudhry et al., 2007). The NS7 mutations KKRAAA, RRAA, R396A and K169A were non-viable, while the recovery of the R395A and K210A mutants was decreased by ~3 and 0.5 $\log_{10}$, respectively (Fig. 3d). The results of the three assays correlated well for most mutations, although it is worth noting that the magnitude to which each mutation affected each activity varied (Fig. 3e). This variability is likely due the inherent properties of each assay. There are a number of notable exceptions, including the mutants K210A and K422A, which appeared to have a minimal impact on virus viability. It is worth noting that the reverse genetic assay used is an endpoint assay and therefore unable to identify mutations that simply slow replication. It is therefore possible that the mutants K210A and K422A may have more subtle phenotypes that may only be apparent in more detailed analysis including *in vivo* virulence studies.

The alanine substitution mutants of arginine residues within motif E (R^{395} and R^{396}) showed decreased levels of RNA binding as well as *de novo* RNA synthesis and viral replication (Figs. 3c and 3d). Collectively, these data suggest a clear role of these restudies in viral RNA binding.

The NTP binding activity of motif F, at the fingers-thumb inter domain, has been described for the HCV RdRp (Lesburg et al., 1999). Mutation of the equivalent amino acids around that motif in MNV-1 K^{169} resulted in a minor effect on RNA binding, but showed a more significant effect on in vitro RNA synthesis and viral recovery by reverse genetics (Fig. 3). These data suggest that this residue may influence RNA synthesis by a mechanism other than RNA binding. When we superimposed our NS7 structure with that of NV polymerase-RNA complex, K^{169} was found to interact with R^{185}, which is bound to CTP. Thus we would propose that K^{169} may function to stabilize NS7-RNA-NTP binding.

In conclusion, we have identified a number of basic amino acids surrounding the norovirus RdRp active site that contribute to viral RNA-binding, RNA synthesis and virus viability. Additional studies will be required to examine the role of the identified amino acids in other functions of the viral polymerase such as VPg guanylylation or the interaction with the viral capsid protein VP1. Furthermore, the potential impact of some of the mutations on viral pathogenesis *in vivo* remains to be determined, but this work provides a framework with which to potentially design rationally attenuated noroviruses.

ACKNOWLEDGEMENT

This study was supported by grants from the Korea Healthcare Technology R&D Project, Ministry for Health, Welfare & Family Affairs (A085119) and Basic Science Research Programs through NRF funded by the Ministry of Education (NRF-2013R1A1A2064940, L.J.-H. and NRF-2016R1A6A3A01012238, K.R.H.). BA was supported by funding from Qassim University, Saudi Arabia, and the work in the lab of IG is supported by the Wellcome Trust (097997/Z/11/Z). K.R.H. was a recipient of postdoctoral fellowship from the BK21+ program. IG is a Wellcome Senior Fellow.

References

Bressanelli S, Tomei L, Rey FA, De Francesco R. Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J Virol* 2002;76:3482-3492.

Bressanelli S, Tomei L, Roussel A, Incitti I, Vitale RL et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc Natl Acad Sci USA* 1999;96:13034-13039.

Chaudhry Y, Skinner MA, Goodfellow IG. Recovery of genetically defined murine norovirus in tissue culture by using a fowlpox virus expressing T7 RNA polymerase. *J Gen Virol* 2007;88:2091–2100.

Daughenbaugh KF, Fraser CS, Hershey JW, Hardy ME. The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment. *EMBO J* 2003;22:2852–2859.

Ettayebi K, Crawford SE, Murakami K, Broughman JR, Karandikar U et al. Replication of human noroviruses in stem cell-derived human enteroids. *Science* 2016;353:1387-1393.

Fukushi S, Kojima S, Takai R, Hoshino FB, Oka T et al. Poly(A)- and primer-independent RNA polymerase of Norovirus. *J Virol* 2004;78:3889–3896.

Fullerton SW, Blaschke M, Coutard B, Gebhardt J, Gorbalenya A et al. Structural and functional characterization of sapovirus RNA-dependent RNA polymerase. *J Virol* 2007;81:1858–1871.

Green KY. Caliciviruses: The Noroviruses. In: Knipe DM, Howley PM (editor). *Fields virology*, 5th ed, Lippincott Williams & Wilkins, Philadelphia, PA; 2007. pp. 949-979.

Han KR, Choi Y, Min BS, Jeong H, Cheon D et al. Murine norovirus-1 $3D^{pol}$ exhibits RNA-dependent RNA polymerase activity and nucleotidylylates on Tyr of the VPg. *J Gen Virol* 2010;91:1713-1722.

Hansen JL, Long AM, Schultz SC. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* 1997;5:1109-1122.

Jones MK, Watanabe M, Zhu S, Graves CL, Keyes LR et al. Enteric bacteria promote human and mouse norovirus infection of B cells. *Science* 2014;346:755-759.

Karst SM, Wobus CE, Lay M, Davidson J, Virgin HW. STAT1-dependent innate immunity to a Norwalk-like virus. *Science* 2003;299:1575-1578.

Karst SM, Wobus CE, Goodfellow IG, Green KY, Virgin HW. Advances in Norovirus Biology. *Cell Host Microbe* 2014;15:668–680.

Lee JH, Alam I, Han KR, Cho S, Shin S et al. Crystal structures of murine norovirus-1 RNA-dependent RNA polymerase. *J Gen Virol* 2011;92:1607-1616.

Lesburg CA, Cable MB, Ferrari E, Hong Z, Mannarino AF et al. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol* 1999;6:937-943.

McFadden N, Bailey D, Carrara G, Benson A, Chaudhry Y et al. Norovirus regulation of the innate immune response and apoptosis occurs via the product of the alternative open reading frame 4. *PLoS Pathog* 2011;7(12):e1002413.

Medvedev A, Viswanathan P, May J, Korba B. Regulation of human norovirus VPg nucleotidylylation by ProPol and nucleoside triphosphate binding by its amino terminal sequence in vitro. *Virology* 2017;503:37-45.

Ng KK, Cherney MM, Vázquez AL, Machín A, Alonso JM et al. Crystal structures of active and inactive conformations of a caliciviral RNA-dependent RNA polymerase. *J Biol Chem* 2002;277:1381-1387.

Ng KK, Pendás-Franco N, Rojo J, Boga JA, Machín A et al. Crystal structure of Norwalk virus polymerase reveals the carboxyl terminus in the active site cleft. *J Biol Chem* 2004;279:16638–16645.

Olspert A, Hosmillo M, Chaudhry Y, Peil L, Truve E et al. Protein-RNA linkage and posttranslational modifications of feline calicivirus and murine norovirus VPg proteins. *PeerJ* 2016;4:e2134.

Rohayem J, Jager K, Robel I, Scheffler U, Temme A et al. Characterization of norovirus $3D^{pol}$ RNA-dependent RNA polymerase activity and initiation of RNA synthesis. *J Gen Virol* 2006;87:2621–2630.

Sosnovtsev SV, Belliot G, Chang KO, Prikhodko VG, Thackray LB et al. Cleavage map and proteolytic processing of the murine norovirus nonstructural polyprotein in infected cells. *J Virol* 2006;80:7816–7831.

Thorne LG, Goodfellow IG. Norovirus gene expression and replication. *J Gen Virol* 2014;95:278–291.

Vázquez AL, Martín Alonso JM, Parra F. Mutation analysis of the GDD sequence motif of a calicivirus RNA-dependent RNA polymerase. *J Virol* 2000;74:3888-3891.

Wobus CE, Karst SM, Thackray LB, Chang KO, Sosnovtsev SV et al. Replication of norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS Biol* 2004;2:2076-2084.

FIGURE LEGENDS

Fig. 1. Preparation of recombinant NS7 proteins based on amino acid sequence alignment.

(A) Amino acids involved in RNA binding. Alignment of amino acid sequences of RdRps from viruses belonging to the family *Caliciviridae* (NCBI Entrez accession numbers: MNV-1, NC008311; Norwalk, AAB50465; RHDV, P27410). Six motifs (A, B, C, D, E, and F) and 3 domains (fingers, blue; palm, yellow; thumb, green) are indicated. Well-conserved basic amino acids and cation-binding Asp residues are indicated in red and cyan, respectively. (B) Purification of recombinant proteins. Wild type and mutant NS7 proteins purified as described previously (Han et al., 2010) were analyzed by 15% SDS-PAGE. Molecular masses of standard proteins are shown in kDa (left).

Fig. 2. Superposition of MNV-1 NS7 and NV polymerase-RNA complex structures.

(A) Overall view of superimposed structures of MNV-1 NS7 (PDB ID 3QID) and NV polymerase-RNA complex (PDB ID 3BSO): front view (left) and side view (right) where a box is indicated to present important basic amino acids. Double-stranded RNA and CTP from NV polymerase-RNA complex are shown in grey. Fingers, palm and thumb domains are shown in blue, yellow and green, respectively, as shown in Fig. 1a. Basic amino acids are shown as stick model in domain colors and the GDD motif of MNV-1 NS7 is shown in cyan. (B) A close-up view of the box in (A), showing double-stranded RNA, CTP (dark grey), and basic amino acids.

Fig. 3. RNA binding affects RNA synthesis and viral replication.

(A) EMSA of wild type and mutant NS7s. ^{32}P-labeled viral RNA probe was incubated with or without increasing concentrations of wild type or mutant NS7 (1, 2, 4, and 6 μM) and analyzed by TBE/acrylamide gel electrophoresis. The gel was dried, and ^{32}P-labeled RNA bands were visualized using a BAS-1500 phosphorimager. Free probe and RNA-NS7 complexes are indicated. Three separate experiments showed similar results. (B) Quantitative analysis of complex formation of 2 μM wild type and mutant NS7s was determined using the Multi Gauge software from single experiment. PSL-BG values of the areas indicated as complex in (A) were quantified and shown as a percentage of the level of free probe (100%). (C) RNA polymerase activities of mutant NS7s. An RdRp assay with wild type, mutant NS7 or without NS7 was incubated for 30 min at 37°C. D^{346} and D^{347} to Ala double mutant also tested as a control (GAA). The reaction was stopped by adding 200 mM EDTA (pH 8.0).

The incorporated [α-^{32}P]GMP was measured using a liquid scintillation counter. Incorporated GMP is shown as a percentage of the level of the reaction of wild type NS7 (100%). Standard deviations of triplicate independent samples are shown as vertical bars. Statistical analysis was performed by a one-way ANOVA (**P<0.01, *P<0.05). (D) Impact of mutations on MNV-1 virus recovery. cDNA plasmids containing either the wild type (WT) MNV-1 cDNA or NS7 mutants under the control of a T7 RNA polymerase promoter were transfected into BSRT-7 cells previously infected with fowl pox expressing T7 RNA polymerase. 48 hr post transfection, the cells were frozen and thawed and the clarified supernatant used to determine the 50% tissue culture infectious doses (TCID50/mL) in murine microglial cells BV2 cell line. Standard deviations of biological triplicate infections are shown as vertical bars. Statistical analysis was performed by a one-way ANOVA (***P<0.001, **P<0.01, NS: not significant).